AD-17660

UNCLASSIFIED

ESTIMATION OF THE PARAMETERS OF A BIVARIATE NORMAL POPULATION FROM TRUNCATED SAMPLES

by

WM. JACKSON HALL

Institute of Statistics, University of North Carolina

Special report to the Office of Naval Research
of work at Chapel Hill under Project NR 042 031
for research in probability and statistics.

Institute of Statistics
Mimeograph Series No. 78

August 1953

Introduction and summary. Maximum likelihood estimates of the parameters of a bivariate normal distribution are obtained for a sample in which only those observations falling in a specific region can be measured, all other observations being called "unmeasured observations". Two cases are treated, the number of unmeasured observations being unknown (Case I) or known (Case II). Explicit expressions are obtained when the region of truncation is a rectangle or an infinite strip. The asymptotic covariance matrix is obtained simultaneously with the solution.

We denote the bivariate normal density function with parameters μ_x, σ_x, μ_y, σ_y, and ρ (sometimes denoted λ_1, λ_2, λ_3, λ_4, λ_5 for convenience) by $\phi(x,y)$. Then, in Case I, the likelihood of a sample of n independent observations all in a region R is

$$\frac{1}{p^n} \prod_{i=1}^{n} \phi(x_i, y_i)$$

where

$$p = \Pr[(x,y) \text{ in } R] = \int_R \phi(x,y)\, dx\, dy \;;$$

and, in Case II, the likelihood of a sample of N independent observations of which n observations occur in R and N-n elsewhere is

$$\binom{N}{n} (1-p)^{N-n} \prod_{i=1}^{n} \phi(x_i, y_i) \ .$$

The partial derivative of the logarithmic likelihood, L , with respect to one of the parameters, say λ, is

$$\frac{\partial L}{\partial \lambda} = - f(n,p) \frac{\partial p}{\partial \lambda} + \sum_{i=1}^{n} \frac{\partial}{\partial \lambda} \log \phi(x_i, y_i)$$

where $f(n,p) = n/p$ in Case I and $f(n,p) = (N-n)/(1-p)$ in Case II. To obtain the maximum likelihood estimates, all five partial derivatives are equated to zero and solved for the unknown parameters.

<u>Iterative solution of the maximum likelihood equations</u>. To solve the five estimating equations simultaneously, we propose a Newton iterative procedure. Choosing an initial trial solution, we approximate the system of equations by a linear system using the linear terms in a Taylor series expansion. Thus

$$0 = \frac{\partial L}{\partial \lambda_j} \cdot \left(\frac{\partial L}{\partial \lambda_j}\right)_{(1)} + \sum_{i=1}^{5} \left(\frac{\partial^2 L}{\partial \lambda_i \partial \lambda_j}\right)_{(1)} (\lambda_i - \lambda_j^{(1)}) \quad (j=1,\ldots,5)$$

where a subscript (1) denotes evaluation at the first trial point and a superscript (1) denotes the first trial value. In matrix notation we have

$$\ell_{(1)} = A_{(1)}\, d^{(1)}$$

where ℓ is the (column) vector with elements $\frac{\partial L}{\partial \lambda_j}$ $(j=1,\ldots,5)$, $d^{(1)}$ is the (column) vector with elements $(\lambda_j - \lambda_j^{(1)})$, and

$$A = (a_{ij}) = \left(-\frac{\partial^2 L}{\partial \lambda_i \partial \lambda_j}\right) .$$

Second trial values are obtained from the first from

$$d^{(1)} = A_{(1)}^{-1}\, \ell_{(1)}$$

(assuming $A_{(1)}$ non-singular), and by substituting these values for the initial ones further estimates are obtained, and so on until stability is reached. (It may not be necessary to recalculate the A matrix at each step if its elements are sufficiently stationary. When it is recalculated, its inverse may be obtained quickly by iteration [1].)

<u>Precision of Estimates</u>. The asymptotic covariance matrix is the inverse of the matrix with elements $E\left(-\frac{\partial^2 L}{\partial \lambda_i \partial \lambda_j}\right)$, and thus may be estimated by A^{-1}. This estimate is obtained simultaneously with the solution of the estimating equations.

<u>Rectangular truncation</u>. Here we develop explicitly the estimating equations for a rectangularly truncated population. Let the region R be a rectangle, bounded by the lines $x = h_1$, $x = h_2$, $y = k_1$, $y = k_2$ ($h_1 < h_2$, $k_1 < k_2$). Then

$$p = \int_{k_1}^{k_2} \int_{h_1}^{h_2} \phi(x,y)\, dx\, dy .$$

Now $\phi(x,y)$ may be expressed as a power series in ρ with Hermite functions as coefficients

$$\phi(x,y) = \frac{1}{\sigma_x \sigma_y} \sum_{\nu=0}^{\infty} \frac{\rho^\nu}{\nu!} G_\nu\left(\frac{x-\mu_x}{\sigma_x}\right) G_\nu\left(\frac{y-\mu_y}{\sigma_y}\right)$$

where

$$G_\nu(t) = (-1)^\nu \frac{1}{\sqrt{2\pi}} \frac{d^\nu}{dt^\nu} e^{-t^2/2} \qquad (\nu = 1,2,\ldots)$$

$$G_0(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2} .$$

(See R. A. Fisher's introduction, pp.xxvi-xxviii, [2].) For negative subscripts, the Hermite functions are defined by the two following relations:

$$G_{-1}(t) = 1 - \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

$$t\, G_\nu(t) = \nu\, G_{\nu-1}(t) + G_{\nu+1}(t) .$$

Since the series in ρ converges uniformly in x and y, and since $\int G_\nu(t)\, dt = - G_{\nu-1}(t)$, we have

$$p = \int_{\eta_1}^{\eta_2} \int_{\xi_1}^{\xi_2} \sum_{\nu=0}^{\infty} \frac{\rho^\nu}{\nu !} G_\nu(x)\, G_\nu(y)\, dx\, dy$$

$$= \sum_{\nu=0}^{\infty} \frac{\rho^\nu}{\nu !} \int_{\xi_1}^{\xi_2} G_\nu(x)\, dx \int_{\eta_1}^{\eta_2} G_\nu(y)\, dy$$

$$= \Sigma_{0,0}$$

where

$$\xi_i = \frac{h_i - \mu_x}{\sigma_x} \qquad \eta_i = \frac{h_i - \mu_y}{\sigma_y} \qquad (i=1,2)$$

and

$$\Sigma_{r,s} = \sum_{\nu=0}^{\infty} \frac{\rho^\nu}{\nu !} \left[G_{\nu+r-1}(\xi_1) - G_{\nu+r-1}(\xi_2) \right] \left[G_{\nu+s-1}(\eta_1) - G_{\nu+s-1}(\eta_2) \right] .$$

After calculating the derivatives of the $\Sigma_{r,s}$ function, we obtain the following derivatives·

$$(1)\quad \left\{\begin{aligned} \frac{\partial p}{\partial \mu_x} &= \frac{1}{\sigma_x}\Sigma_{1,0} \\ \frac{\partial p}{\partial \sigma_x} &= \frac{1}{\sigma_x}(\rho\Sigma_{1,1} + \Sigma_{2,0}) \\ \frac{\partial p}{\partial \rho} &= \Sigma_{1,1}\ ; \end{aligned}\right.$$

$$(2)\quad \left\{\begin{aligned} \frac{\partial^2 p}{\partial \mu_x^2} &= \frac{1}{\sigma_x^2}\Sigma_{2,0} \\ \frac{\partial^2 p}{\partial \mu_x \partial \sigma_x} &= \frac{1}{\sigma_x^2}(\rho\Sigma_{2,1} + \Sigma_{3,0}) \\ \frac{\partial^2 p}{\partial \mu_x \partial \mu_y} &= \frac{1}{\sigma_x\sigma_y}\Sigma_{1,1} \\ \frac{\partial^2 p}{\partial \mu_x \partial \sigma_y} &= \frac{1}{\sigma_x\sigma_y}(\rho\Sigma_{2,1} + \Sigma_{1,2}) \\ \frac{\partial^2 p}{\partial \mu_x \partial \rho} &= \frac{1}{\sigma_x}\Sigma_{2,1} \end{aligned}\right.$$

$$\left\{\begin{aligned}
\frac{\partial^2 p}{\partial \sigma_x^2} &= \frac{1}{\sigma_x^2}\left(\Sigma_{4,0} + 2\rho\Sigma_{3,1} + \rho^2\Sigma_{2,2} + \Sigma_{2,0}\right) \\
\frac{\partial^2 p}{\partial \sigma_x \partial \sigma_y} &= \frac{1}{\sigma_x\sigma_y}\left[(1+\rho^2)\Sigma_{2,2} + \rho(\Sigma_{3,1} + \Sigma_{1,3} + \Sigma_{1,1})\right] \\
\frac{\partial^2 p}{\partial \sigma_x \partial \rho} &= \frac{1}{\sigma_x}\left(\rho\Sigma_{2,2} + \Sigma_{3,1}\right) \\
\frac{\partial^2 p}{\partial \rho^2} &= \Sigma_{2,2} ;
\end{aligned}\right.$$

all others being obtained by symmetry. (The interchange of x and y requires the interchange of the order of the subscripts on the $\Sigma_{r,s}$ functions.)

In Case I,

$$\frac{\partial}{\partial p} f(n,p) = -n/p^2 \quad . \tag{3}$$

and in Case II,

$$\frac{\partial}{\partial p} f(n,p) = (N-n)/(1-p)^2 \; . \tag{4}$$

Calculating the derivatives of log $\emptyset(x,y)$, we find

$$
\text{(5)}\quad \left\{
\begin{aligned}
\sum_{i=1}^{n} \frac{\partial}{\partial \mu_x} \log \phi(x_i, y_i) &= -\frac{n}{\sigma_x(1-\rho^2)} (\rho\, m_{01} - m_{10}) \\
\sum \frac{\partial}{\partial \sigma_x} \log \phi &= -\frac{n}{\sigma_x(1-\rho^2)} (1 - \rho^2 + \rho\, m_{11} - m_{20}) \\
\sum \frac{\partial}{\partial \rho} \log \phi &= -\frac{n}{(1-\rho^2)^2} \left[\rho(m_{20} + m_{02} - 1 + \rho^2) - (1 + \rho^2) m_{11} \right];
\end{aligned}
\right.
$$

$$
\text{(6)}\quad \left\{
\begin{aligned}
\sum_{i=1}^{n} \frac{\partial^2}{\partial \mu_x^2} \log \phi(x_i, y_i) &= -\frac{n}{\sigma_x^2(1-\rho^2)} \\
\sum \frac{\partial^2}{\partial \mu_x \partial \sigma_x} \log \phi &= -\frac{n}{\sigma_x^2(1-\rho^2)} (2 m_{10} - \rho\, m_{01}) \\
\sum \frac{\partial^2}{\partial \mu_x \partial \mu_y} \log \phi &= \frac{n \rho}{\sigma_x \sigma_y (1-\rho^2)} \\
\sum \frac{\partial^2}{\partial \mu_x \partial \sigma_y} \log \phi &= \frac{n \rho\, m_{01}}{\sigma_x \sigma_y (1-\rho^2)} \\
\sum \frac{\partial^2}{\partial \mu_x \partial \rho} \log \phi &= -\frac{n}{\sigma_x(1-\rho^2)^2} \left[(1 + \rho^2) m_{01} - 2\rho\, m_{10} \right]
\end{aligned}
\right.
$$

$$
\left\{
\begin{aligned}
\Sigma \frac{\partial^2}{\partial \sigma_x^2} \log \phi &= -\frac{n}{\sigma_x^2(1-\rho^2)} (3\, m_{20} - 2\rho\, m_{11} - 1 + \rho^2) \\
\Sigma \frac{\partial^2}{\partial\sigma_x \partial\sigma_y} \log \phi &= \frac{n\rho\, m_{11}}{\sigma_x \sigma_y (1-\rho^2)} \\
\Sigma \frac{\partial^2}{\partial\sigma_x \partial\rho} \log \phi &= -\frac{n}{\sigma_x(1-\rho^2)^2} \left[(1+\rho^2)\, m_{11} - 2\rho\, m_{20} \right] \\
\Sigma \frac{\partial^2}{\partial\rho^2} \log \phi &= -\frac{n}{(1-\rho^2)^3} \left[(1+3\rho^2)(m_{20}+m_{02}) \right. \\
&\qquad \left. - 2\rho\,(3+\rho^2)\, m_{11} - (1-\rho^4) \right] ;
\end{aligned}
\right.
$$

all others being obtained by symmetry; we have denoted

$$
m_{rs} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{x_i - \mu_x}{\sigma_x}\right)^r \left(\frac{y_i - \mu_y}{\sigma_y}\right)^s \qquad (r,s = 0,1,2) .
$$

Using (1) and (5), we may now calculate the elements of the ℓ vector:

$$
\frac{\partial L}{\partial \lambda} = -f(n,p) \frac{\partial \upsilon}{\partial \lambda} + \sum_{i=1}^{n} \frac{\partial}{\partial \lambda} \log \phi(x_i, y_i) \quad ;
$$

and using (1), (2), (3), (4), and (6), we may calculate the elements of the A matrix:

$$(8) \quad -\frac{\partial^2 L}{\partial\lambda_j \partial\lambda_k} = f(n,p)\frac{\partial^2 p}{\partial\lambda_j\partial\lambda_k} + \frac{\partial f}{\partial p}\frac{\partial p}{\partial\lambda_j}\frac{\partial p}{\partial\lambda_k} - \sum_{i=1}^{n}\frac{\partial^2}{\partial\lambda_j\partial\lambda_k}\log \phi(x_i, y_i) .$$

(The required $\Sigma_{r,s}$ functions may be computed from tables of the tetrachoric functions [3], using the relation $\sqrt{\nu!}\,\tau_\nu(t) = G_{\nu-1}(t)$.)

The estimates for truncation over an infinite quarter-plane may be obtained by letting one of the h's and one of the k's in the discussion above go to $\pm\infty$.

Initial estimates may be obtained by approximating the region R by an infinite strip and using the linear truncation estimation method which follows.

Linear truncation. We shall consider independently the estimation problem when the region R is an infinite strip, bounded by $x = h_1$ and $x = h_2$ ($h_1 < h_2$). Here, we shall distinguish three cases: (I) the number of unmeasured observations is unknown, (II) the numbers of unmeasured observations in each truncated half-plane are known (say n_1 observations in $R_1 = [(x,y): x < h_1]$ and n_2 in $R_2 = [(x,y): x > h_2]$, $n_1 + n_2 = N-n$), and (III) only the total number N-n of unmeasured observations is known.

Now the marginal distribution of x is independent of μ_y, σ_y, and ρ, and is simply a truncated univariate normal distribution. Thus, in all three cases, μ_x and σ_x may be estimated by the methods of A. C. Cohen [4] for truncated univariate normal distributions, the three cases above corresponding to the three cases enumerated by Cohen.

The likelihood functions in the three cases are:

$$
(9)\quad \begin{cases} \text{(I)} & \dfrac{1}{p^n} \prod_{i=1}^{n} \phi(x_i, y_i) \\[2ex] \text{(II)} & k_1\, p_1^{\,n_1}\, p_2^{\,n_2} \prod_{i=1}^{n} \phi(x_i, y_i) \\[2ex] \text{(III)} & k_2 (1-p)^{N-n} \prod_{i=1}^{n} \phi(x_i, y_i) \end{cases}
$$

where

$$p = \Pr[(x,y) \text{ in } R] = \Phi\left(\frac{h_2-\mu_x}{\sigma_x}\right) - \Phi\left(\frac{h_1-\mu_x}{\sigma_x}\right)$$

$$p_1 = \Pr[(x,y) \text{ in } R_1] = \Phi\left(\frac{h_1-\mu_x}{\sigma_x}\right)$$

$$p_2 = \Pr[(x,y) \text{ in } R_2] = 1 - \Phi\left(\frac{h_2-\mu_x}{\sigma_x}\right)$$

and

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt \; ;$$

k_1 and k_2 are constants. Since p, p_1, and p_2 are all independent of μ_y, σ_y, and ρ, the maximum likelihood equations for these three parameters will be the same in all three cases; namely, from (5) and (9),

$$\frac{\partial L}{\partial \mu_y} = -\frac{n}{\sigma_y(1-\rho^2)} (\rho\, m_{10} - m_{01}) = 0$$

$$\frac{\partial L}{\partial \sigma_y} = -\frac{n}{\sigma_y(1-\rho^2)}(1-\rho^2+\rho\, m_{11}-m_{02}) = 0$$

$$\frac{\partial L}{\partial \rho} = -\frac{n}{(1-\rho^2)^2}\left[\rho\,(m_{20}+m_{02}-1+\rho^2)-(1+\rho^2)\,m_{11}\right] = 0$$

where L denotes the logarithmic likelihood, m_{rs} being defined by (7). Having obtained estimates of μ_x and σ_x by Cohen's method, these three equations may be solved for estimates of μ_y, σ_y, and ρ, yielding (after some algebraic manipulation):

$$\mu_y = \frac{(m_{11}' - m_{01}'\,\mu_x)(m_{10}' - \mu_x) - m_{01}'(m_{20}' - 2\,m_{10}'\,\mu_x + \mu_x^2)}{(m_{10}' - \mu_x)^2 - (m_{20}' - 2\,m_{10}'\,\mu_x + \mu_x^2)}$$

(10) $$\sigma_y^2 = m_{02}' - 2\,m_{01}'\,\mu_y + \mu_y^2 + \left(\frac{m_{01}' - \mu_y}{m_{10}' - \mu_x}\right)^2 \left[\sigma_x^2 - (m_{20}' - 2\,m_{10}'\,\mu_x + \mu_x^2)\right]$$

(11) $$\rho = \frac{\sigma_x}{\sigma_y}\,\frac{m_{01}' - \mu_y}{m_{10}' - \mu_x}$$

where m_{rs}' denotes sample moments about the origin. If $\mu_x = m_{10}'$, then we substitute

$$\frac{m_{11}' - m_{10}'\,m_{01}'}{m_{20}' - m_{10}'^2} \quad \text{for} \quad \frac{m_{01}' - \mu_y}{m_{10}' - \mu_x}$$

in equations (10) and (11).

Since the estimates of μ_x and σ_x are independent of μ_y, σ_y, and ρ, we find that the A matrix is now the direct sum of two sub-matrices, and likewise for its inverse. The first inverse sub-matrix (corresponding to μ_x and σ_x) is given by Cohen [4], the second may be obtained by inverting the matrix whose elements are given by (8).

Acknowledgement. Acknowledgement is hereby made to Professor H. Hotelling for proposing this problem and for helpful suggestions, and to Professor S. N. Roy for reading the manuscript.

REFERENCES

[1] Harold Hotelling, "Practical problems of matrix calculation", Berkeley Symposium on Mathematical Statistics and Probability, University of California Press (1949), pp. 275-93.

[2] British Association for the Advancement of Science, Mathematical Tables, Vol. I, Cambridge University Press.

[3] Karl Pearson, Tables for Statisticians and Biometricians, Part II, Cambridge University Press (1931), pp. 74-7.

[4] A. C. Cohen, Jr., "Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples", Annals of Mathematical Statistics, Vol. XXI (1950), pp. 557-69.